Open Access

Phytopathogenomics and Disease Control*., *2024, 3(1):31-39
ISSN (Online):2957-5842; ISSN (Print):2957-5834
DOI: https://doi.org/10.22194/Pdc/3.1032
https://societyfia.org/journal/PDC

Society for Innovative Agriculture

Genome-Wide in Silico Analysis of the Heat Shock Protein (HSP) Family in Vigna Species: Insights into Molecular Mechanisms of Stress Response

Misha Shahzadi, Zakia Habib and Siddra Ijaz

Centre of Agricultural Biochemistry and Biotechnology, University of Agriculture, Faisalabad-38040, Pakistan
*Corresponding author's e-mail: zakianoor92@gmail.com

Heat shock proteins, HSPs, are essential for plants to develop defense systems against various abiotic challenges. *Vigna* species are widely grown cash crops in the world. However, some unfavorable environmental conditions hinder the growth and production of these crops. It is therefore essential to investigate the factors influencing their growth, development, and tolerance to abiotic stress to develop stress-tolerant varieties. Consequently, a comprehensive genome-wide identification study was conducted to identify the Heat Shock Factor (HSF) genes in three *Vigna* species genomes employing computational tools. In the current study, 43 non-redundant HSP proteins were identified in three Vigna species including 17 in *Vigna angularis*, 22 in *Vigna radiata*, and 4 non-redundant in *V. unguculata*. Moreover, subcellular locations were predicted, and all HSPs were found in the nucleus. The theoretical isoelectric point predicted that most of the HSP proteins were acidic. The aliphatic index of the HSP proteins demonstrated that all proteins are highly thermostable. Moreover, the instability index revealed two proteins VangHSP_9 and VradHSP_9, considered stable while the rest of the HSP proteins were unstable. The phylogenetic analysis categorized HSP genes into two major clusters and gene structure analysis displayed that all HSP genes were intronless. This study gives novel insight into the functional characterization of HSPs in *Vigna* species for plant breeding programs.

Keywords: Vigna species, HSPs, In-silico analysis, biotic and abiotic stresses, Vigna species, Abiotic stress, Isoelectric point, Thermostability, Instability index.

INTRODUCTION

Heat shock factors (Hsfs) are a category of transcription factors that are normally dormant but can get activated rapidly in reaction to high temperatures. They are important regulators of growth responses (Zhang *et al.*, 2021). All living things, from bacteria to humans, are members of the highly conserved and ancient protein family known as heat shock proteins (Juneja *et al.*, 2023; Mondal *et al.*, 2023). These proteins are involved in several cellular activities that are vital for preserving cellular homeostasis. Although their activation in response to high temperatures was the initial reason for their discovery, these proteins serve vital functions beyond thermal stress (Zhou *et al.*, 2022). Based on their molecular weights, HSPs are classified into several families, such as small heat shock proteins (sHSPs), HSP60, HSP70, HSP90, and HSP100. Every family performs specific functions for the cell and adapts to a range of stressful circumstances, not just heat stress (Wu *et al.*, 2022).

Beans are a substantial source of human and animal diets having high nutritional value. Beans belong to the *Vigna* genus having flowering plants in the *Fabaceae* family (Habib, 2023). The genus *Vigna* consists of about 200 species having pantropical distribution, native to tropical regions worldwide. Several species of *Vigna* genus are considered as potential protein sources for millions of people in developing countries (Atteh and Adeyeye, 2022). Many *Vigna* species like mung bean (*V. radiate*), cowpea (*V. unguiculata*), mat bean (*V. aconitifolia*), urd bean (*V. mungo*), rice bean (*V. umbellate*), bambara groundnuts (*V. subterranea*) and adzuki bean (*V. angularis*) are keyfood staples as standard diet. These *Vigna* species exhibited considerable economic importance in recent decades (Karami *et al.*, 2022; Wang *et al.*, 2022).

These valuable *Vigna* species are vulnerable to many environmental stresses as well as biological stresses (Singh *et al.*, 2022). These are the major factors that cause adverse effects on the productivity of beans. Environmental stress includes extreme temperature (cold or heat), drought, salinity,

Shahzadi, M., Z. Habib and S. Ijaz. 2024. Genome-Wide in Silico Analysis of the Heat Shock Protein (HSP) Family in Vigna Species: Insights into Molecular Mechanisms of Stress Response. Phytopathogenomics and Disease Control 3:31-39.
[Received 15 Dec 2023; Accepted 13 Mar 2024; Published 4Apr 2024]

water-loggings, and oxidative stresses (Bhardwaj *et al.*, 2023). Other abiotic stresses include ultraviolet or ionizing radiations, metal pollutants, insecticide residue, prolonged rainy conditions, and many their mechanical factors like wind and water pressure (Tutlani *et al.*, 2022). Biological factors involve many living organisms that cause particular diseases in beans. Gene that responds to these stresses play a vital role in plant physiology (Ahanger *et al.*, 2017).

In silico Genome-wide analysis of these genes will provide an opportunity to develop resistant-linked markers and genetic maps and phylogenetic relationship for significantly studying the plant genome (Chen *et al.*, 2021; Ijaz *et al.*, 2023b). However, these genes are controlled by several transcription factors that regulate their functioning in stressfull environment. The genome wide analysis of the genes linked to HSP family has been conducted in many plants that provided characterization and evolutionary background of these genes (Ren *et al.*, 2023; Shamshad *et al.*, 2023). But in the case of Vigna species, studies related to HSP factors have not yet been conducted. In current study, in silico characterization of Heat Shock Protein (HSP) family has been investigated throughout the whole genome of Vigna species.

MATERIALS AND METHODS

The Heat Shock Protein (HSP) represents a significant transcription factor uniquely found in plants. These HSP factors are instrumental in governing various cellular processes in plants, including growth, differentiation, development, and their reactions to both biotic and abiotic stresses. Although genome wide analysis of HSP protein family has been explored in many crop species but in case of *Vigna* species, no relevant study has been emerged yet. The present study was conducted to identify the HSP genes, intron-exon architecture, and motif distribution as well as the comparative evolutionary relationship of non-redundant HSP proteins among Vigna species in reference to *Arabidopsis* using computational biology.

***Sequence retrieval*:** A database specified to plant transcription factors, known as PlantTFDB (Plant Transcription Factor database) version 5.0 (http://planttfdb.gao-lab.org/) was used for obtaining HSP proteins sequences and CDS in three Vigna species (*Vigna angularis*, *Vigna radiate*, and *Vigna unguiculata*). However, genomic sequences of the HSP protein of three Vigna species and Arabidopsis were searched from NCBI (National Center for Biotechnology Information) (https://www.ncbi.nlm.nih.gov/) database.

***Conserved Domain/ motif analysis*:** Online web server, InterProScan was used to identify conserved HSP domains among collected HSP protein sequences. Retrieved sequences of HSP proteins were subjected to the ScanProsite tool (https://prosite.expasy.org/scanprosite/). The conserved motifs within these retrieved protein sequences were identified using MEME (Multiple Expression motifs for Motif Elicitation) search tool (http://meme-suite.org/tools/meme). In MEME, a maximum number of motifs was set at 15 in three *Vigna* species.

***Physicochemical properties*:** Online protein servers, ProtParam (https/web.expasy.org/protparam/) and Plant1FDB 5.0 were used for collecting information about amino acid composition, isoelectric point *(pl)*, molecular weight, molecular formula, instability index (10), aliphatic index and grand average of hydropathicity (GRAVY) of HSP proteins along with chromosomal location of the HSP proteins of Vigna species.

***Protein subcellular localization*:** The prediction of subcellular localization of HSP protein in three Vigna species was performed using CELLO v.2.5 (https//cello.life.nctu.edu.tw/). This online server provides location of HSP proteins within cell based on SVMs (support vector machines) classification system.

***Nuclear localization signals (NLS)*:** NLStradamus (http://www.moseslab.csb.utoronto.ca/NIStradamus/) was used for predicting NLS regions in investigated HSP proteins of *Vigna* species. This NLS signal prediction was based on hidden markov model.

***Gene Structure analysis*:** In gene structure display analysis, gene sequences were retrieved from their respective selected HSP proteins. The genes structure with intron-exon architecture of retrieved HSP genes was explored through Gene Structure Display Server 2.0 (GDS) (https://gsds.cbi.pku.edu.cn/).

***Phylogenetic analysis*:** The HSP sequences from three Vigna species were aligned using CLUSTALW program (https://www.genome.jp/tools-bin/clustalw). The Phylogenetic trees were generated using the Molecular Evolutionary Genetics Analysis v. 7.0 (MEGA7) software package (Kumar *et al*. 2016) under the neighbor-joining algorithm

RESULTS

In a genome-wide investigation, genes linked to specific cellular processes within plant species are characterized. Although genome-wide analyses of the *HSP* family have been investigated in many crop species, no studies specifically addressing *Vigna* species have yet to be published. With the use of computational biology, the current work aimed to discover the *HSP* genes, their intron-exon architecture, motif distribution, and comparative evolutionary relationships of non-redundant HSP among *Vigna* species. Therefore, the goal of this research work was *"In silico* Genome-wide probing and analysis of HSP family in *Vigna* species (*Vigna radiata, Vigna angularis, Vigna unguiculata*) for the characterization and function prediction".

***Genome-wide identification of Vigna proteins*:** In genome-wide analysis, a total of 43 non-redundant HSP proteins in *Vigna* species (17 in *Vigna angularis*, 22 in *Vigna radiata*,

Table 1. Conserved motif in identified HSP proteins of *Vigna* species, predicted by MEME suite.

Motif	Protein motif of *V. angularis*	Protein motif of *V. radiata*	Protein motif of *V. ungucuilata*
1	TDDIVSWSDDGNSFVVWDPAEFARDLLPRYFKHNNFSSF	DDPSTBDIVSWSEDGNSFVVWBPPEFARDLLPKYFKHNNFSSF	FIVWNPPEFARHLLPKYFKHNNFSSFIRQLNTYGFRKVDPDRWEFANEKF
2	RQLNTYGFRKVDPDRWEFANEGFLRGQKHLLKEIRRRKSPS	GFRKVDPDRWEFANEGFLRGQKHLLKNIRRRKPAHSHSAQQ	PFLLKTYDLVDDPSTDDIISW
3	VLMQELVKLRQQQQNTRNQLQEMEQRLQGTEQKQQQMMSFLAKALQNPAF	EIEKLKRDKQVLMQELVRLRQQQQTTKNQLQEMEQRLZGTEQKQQQMMSF	RGQKHLLKEIQRRKP
4	LQESGPPPFLTKTYDLVDDPS	SGPPPFLTKTYDLV	NANGIGGGQADAGQGGEQGCR
5	GACVEVGKFGLEGEVERLKRD	RQLNTY	FQMQID
6	ENERLKKENEVLSKELAEMKSLCNEJLAL	LAKALQNPAFLQQLV	RLDATEKRQKQLLPFLSKALZDPQVVSH
7	AASDGQIVKYQPDWNEAAKAS	LLEENERLKKENLVLQSELEEMKSLCNEJLDLLSNY	KKRRLC
8	NPHFWDEILRTPVPEDIETNGAEVFNENELQPTENGWEKSQRMDHLTEQM	QSVGACVEVGKFGLD	MPLSAI
9	QLLQQKEKRKELEEALSKKRR	GVNDVFWEQFLTERPGC	TAPATVVVP
10	FRKIDPDRWEFANEGFLRGQKHLLRNIRRRK	DGQIVKYQPDINNAAKS	GGDCGE
11	PAIPVSQADEIIPDLPPIPEIVAGNILDIPZENYMA	EEAANKKRRLPQEGI	RLHPPG
12	DAIWEDLLNZDLVAGDPEEEVIIGDCSQIDVPVEDLVADPDEW	PHFWDEILRTPVPEDIETNGPEVFNENEVQPTENGWEKSQRMDHLTEQMG	HQHTAK
13	HPGPGDVKAEPLDCZRHGENREAPWLNQCRIANZREC	PKLFGVSLKGKKGR	ADSQRS
14	SPKLFGVALG	KFWWNIRNVNNPPEPMGH	IFGEDAAE
15	YJIGDASSPSSEMDRGGSLGWSSGVTLKEVPPAKVQSSHVPAAAGTQGH	GVTLKEVPPAKVQSSHVPAATGTQGHGPPTGKPEIPSVPQVVACEEVTK	AAPAAGGG

Table 2. Physicochemical properties of identified V*igna angularis* HSP protein family.

Sr. #	Sequence code	AA	MW (Da)	*pI*	Molecular formula	II	Aliphatic Index	GRAVY
1	VangHSP_01	449	50913.44	6.34	$C_{2251}H_{3512}N_{642}O_{679}S_{15}$	46.21	78.82	-0.441
2	VangHSP_02	233	26459.06	5.98	$C_{1177}H_{1843}N_{321}O_{355}S_{9}$	56.23	85.36	-0.375
3	VangHSP_03	338	38571.38	4.99	$C_{1692}H_{2656}N_{472}O_{532}S_{14}$	61.17	78.43	-0.554
4	VangHSP_04	363	41747.43	5.21	$C_{1831}H_{2843}N_{525}O_{578}S_{9}$	69.73	61.98	-0.930
5	VangHSP_05	347	37656.84	5.78	$C_{1624}H_{2573}N_{479}O_{531}S_{11}$	54.31	70.00	-0.574
6	VangHSP_06	211	24488.73	7.01	$C_{1066}H_{1675}N_{307}O_{328}S_{14}$	58.07	65.17	-0.768
7	VangHSP_07	376	42242.11	5.00	$C_{1849}H_{2897}N_{525}O_{590}S_{10}$	62.62	76.97	-0.567
8	VangHSP_08	312	35081.54	8.13	$C_{1524}H_{2430}N_{446}O_{479}S_{13}$	58.58	73.43	-0.575
9	VangHSP_09	402	46029.74	6.37	$C_{2060}H_{3251}N_{565}O_{603}S_{14}$	39.27	87.44	-0.454
10	VangHSP_10	356	41442.05	5.96	$C_{1831}H_{2873}N_{517}O_{550}S_{16}$	54.82	72.81	-0.712
11	VangHSP_11	456	51208.34	4.99	$C_{2254}H_{3533}N_{627}O_{713}S_{12}$	58.09	74.98	-0.611
12	VangHSP_12	206	24048.46	8.99	$C_{1061}H_{1683}N_{301}O_{317}S_{10}$	42.12	72.33	-0.660
13	VangHSP_13	262	30775.96	6.27	$C_{1355}H_{2143}N_{391}O_{407}S_{11}$	61.67	75.08	-0.800
14	VangHSP_14	218	24824.39	8.46	$C_{1090}H_{1740}N_{308}O_{330}S_{12}$	43.62	73.30	-0.628
15	VangHSP_15	263	30081.31	9.50	$C_{1347}H_{2117}N_{373}O_{395}S_{7}$	40.51	77.79	-0.490
16	VangHSP_16	370	42423.88	4.98	$C_{1878}H_{2932}N_{506}O_{584}S_{15}$	51.94	76.59	-0.656
17	VangHSP_17	482	53252.26	4.93	$C_{2316}H_{3623}N_{655}O_{752}S_{18}$	58.19	65.73	-0.636

and 4 *HSP* in *Vigna ungucuilata*) using the Plant Transcription Factor Database (PlantTFDB) and National Centre for Biotechnology Information (NCBI).

***Conserved Domain/ motif analysis*:** The conserved HSP domains were confirmed in 43 non-redundant HSP proteins in *Vigna* species in BatchCD search and ScanProsite. Moreover, fifteen (15) conserved motifs with distinct architecture were identified in the *Vigna* species MEME search tool (Table 1, Figure 1, 2 & 3).

***Physicochemical properties of HSP proteins*:** The physicochemical properties including number of amino acids, molecular weight, isoelectric point, molecular formula, aliphatic index, GRAVY, and instability index of HSP proteins of three *Vigna* species have been explored (Table 2). In *V. angularis*, Lecusine (Leu) was found as a highly abundant amino acid followed by glutamic acid (Glu), valine (Val), aspartic acid (Asp), and arginine (Arg) and thrionine (Thr). In V angularis, the length of amino acids for HSP proteins was varied from 206 for VangHSP_12 to 482 for VangHSP_17. The average length of amino acids was 308.27. The maximum molecular weight (Da) for HSP proteins in VangHSP_17 was 53252.26 and the minimum molecular weight (Da) in VangHSP_12 was 240448.46. The average molecular weight (Da) for HSP protein was 35264.67. The theoretical isoelectric point ranged from 4.93 for VangHSP_17 to 9.50 for VangHSP_15. Most of the VangHSP proteins were in basic range of pI value with an average of 6.40. The instability index is basically a stability

measure of proteins and if the instability index is less than 40 than protein considered to be stable. The instability index (II) for VangHSP_9 was 39.27, considered as stable while the rest of the HSP proteins were unstable. Aliphatic index is the relative volume covered by aliphatic amino acids in side chain. The aliphatic index ranged from 61.98 for VangHSP_4 to 85.36 for VangHSP_2 with an average of 74.48. GRAVY index is the measure of hydrophathy (hydrophilic-positive GRAVY and hydrophobic-negative GRAVY) value for peptides. The GRAVY value of VangHSPs was in negative range 4.2.

In Vigna radiata, arginine (Arg) was found as highly abundant amino acid followed by aspartic acid (Asp), Glutamine (Glu), serine (Ser), and valine (Val). In *V. radiata*, the length of amino acids for HSP proteins was varied from 147 for VradHSPs_21 to 517for VradHSP_14. The average length of amino acids was 356. The maximum molecular weight (Da) for HSP proteins in VradHSP_14 was 60080.09 and the minimum molecular weight (Da) in VradHSP_10 was 20803.81.The average molecular weight (Da) for protein was 30807.79. The theoretical isoelectric point ranged from 4.69 for VradHSP_11 to 9.30 for VradHSP_12. Most of the VradiHSP proteins were in basic range of pl value with an average of 6.172. The instability index for VradHSP_9 was 32.93 considered as stable while the rest of the HSP proteins were unstable. The aliphatic index ranged from 61.41 for VradHSP_4 to 79.69 for VradHSP_16 with an average of 73.70, whereas the average GRAVY value was -0.68 (Table 3).

In *Vigna ungucuilata*, serine (Ser) was found as highly abundant amino acid followed by leucine (Leu) and thrionine (Thr), asparginine (Asn), and glutamine (Glu). In V ungucuilata, the length of amino acids for HSP proteins was varied from 210 for VunHSP_10 to 324 for VunHSP_1. The average length of amino acids was 269.25. The maximum molecular weight (Da) for HSP proteins in VunHSP_2 was 35574.13 and the minimum molecular weight (Da) in VunHSP_11 was 24347.62. The average molecular weight (Da) for *V. ungucuilata* HSP proteins was 30310.48. Their theoretical pI values were ranged from 5.96 for VunHSP_2 to 8.89 for VunHSP_4. Most of the VunHSP proteins were in basic range of pI value with an average of 6.89. In *V. ungucuilata*, all HSP proteins were unstable with more than 40 instability index. The aliphatic index ranged from 65.43 for VunHSP_3 to 73.62 for VunDOF_2 with an average of 69.36. The average GRAVY value was -0.17 (Table 4).

***Subcellular localization and nuclear localization signal prediction in HSP*:** Subcellular localization is the location of

Table 3. Physicochemical properties of identified V*igna radiata* HSP protein family.

Sr. #	Sequence code	AA	MW (Da)	*PI*	Molecular formula	II	Aliphatic index	GRAVY
1	VradHSP_01	402	45594.61	5.11	$C_{1990}H_{3088}N_{568}O_{638}S_{13}$	52.42	68.91	-0.723
2	VradHSP_01	448	50604.39	5.25	$C_{2209}H_{3432}N_{624}O_{706}S_{18}$	60.86	71.65	-0.574
3	VradHSP_01	363	41808.65	5.26	$C_{1833}H_{2846}N_{524}O_{575}S_{12}$	68.39	61.74	-0.893
4	VradHSP_01	192	22103.72	7.70	$C_{958}H_{1517}N_{283}O_{305}S_{7}$	55.38	61.41	-0.886
5	VradHSP_01	462	52517.34	6.52	$C_{2331}H_{3638}N_{664}O_{696}S_{13}$	51.56	79.50	-0.463
6	VradHSP_01	258	29788.91	6.86	$C_{1326}H_{2069}N_{369}O_{391}S_{11}$	64.24	79.30	-0.608
7	VradHSP_01	230	25992.24	6.64	$C_{1152}H_{1789}N_{313}O_{357}S_{8}$	50.58	79.17	-0.437
8	VradHSP_01	506	56016.35	5.13	$C_{2427}H_{3810}N_{698}O_{788}S_{20}$	57.53	67.61	-0.656
9	VradHSP_01	253	28030.30	6.46	$C_{1217}H_{1921}N_{351}O_{392}S_{9}$	32.93	64.35	-0.699
10	VradHSP_01	175	20803.81	7.60	$C_{917}H_{1436}N_{254}O_{273}S_{13}$	67.77	63.54	-0.782
11	VradHSP_01	340	39448.00	4.69	$C_{1724}H_{2700}N_{480}O_{557}S_{12}$	59.22	75.06	-0.843
12	VradHSP_01	186	22156.39	9.30	$C_{978}H_{1561}N_{279}O_{290}S_{9}$	58.38	69.62	-0.768
13	VradHSP_01	452	51023.06	4.94	$C_{2249}H_{3514}N_{624}O_{711}S_{11}$	59.68	74.58	-0.648
14	VradHSP_01	517	60080.09	8.24	$C_{2644}H_{4152}N_{774}O_{787}S_{22}$	39.68	72.24	-0.685
15	VradHSP_01	312	35043.49	6.68	$C_{1522}H_{2428}N_{444}O_{480}S_{13}$	60.70	75.00	-0.569
16	VradHSP_01	358	41182.77	5.76	$C_{1831}H_{2868}N_{508}O_{548}S_{13}$	45.40	79.69	-0.563
17	VradHSP_01	498	56235.49	5.61	$C_{2466}H_{3823}N_{701}O_{783}S_{13}$	57.13	65.66	-0.715
18	VradHSP_01	337	38210.31	4.73	$C_{1669}H_{2584}N_{462}O_{546}S_{11}$	51.80	68.78	-0.692
19	VradHSP_01	465	51741.83	5.45	$C_{2250}H_{3556}N_{636}O_{730}S_{17}$	53.96	74.34	-0.552
20	VradHSP_01	468	52097.53	5.30	$C_{2306}H_{3589}N_{635}O_{715}S_{14}$	54.09	74.44	-0.402
21	VradHSP_01	147	16605.79	7.67	$C_{731}H_{1137}N_{205}O_{222}S_{8}$	43.14	62.24	-0.634
22	VradHSP_01	483	53313.33	4.90	$C_{2322}H_{3630}N_{656}O_{752}S_{17}$	58.28	67.41	-0.636

Table 4. Physicochemical properties of identified *Vigna ungucuilata* HSP protein family.

Sr. #	Sequence code	AA	MW (Da)	*PI*	Molecular formula	II	Aliphatic index	GRAVY
1	Vun002051	324	35364.72	6.13	$C_{1574}H_{2440}N_{438}O_{480}S_{12}$	57.06	71.11	-0.559
2	Vun002151	318	35574.13	5.96	$C_{1574}H_{2450}N_{438}O_{480}S_{12}$	55.65	73.62	-0.508
3	Vun004802	210	24347.62	6.61	$C_{1065}H_{1676}N_{300}O_{328}S_{13}$	59.91	65.43	-0.762
4	Vun009579	225	25955.46	8.89	$C_{1574}H_{2450}N_{439}O_{480}S_{12}$	61.57	67.29	-0.763

proteins that are present in plant cells. In general, all HSP proteins are present in the nucleus (Kushwahan *et al*, 2010; Fang *et al*, 2020). As a result, three *Vigna* species HSP subcellular locations have also been predicted. The predicted localization through software program (DeepLoc-1.0) indicated that all 22 *V. radiata* HSPs were found in the nucleus, while all 17 HSPs of *V. angularis* were also found in nucleus. Further, 4 *V. unguiculata* were also found in the nucleus. Moreover, 4 NLSs were predicted in V. angularis HSPs (VangHSP_03, VangHSP_04, VangHSP_07, and VangHSP_12), two NSLs were predicted in *V. radiata* including VradHSP_03 and VradHSP_08. In contrast, only one NLS was predicted in *V. ungucuilata* (VunHSP_02) (Table 5, 6 & 7).

Table 5. Nuclear localization signal in *Vigna angularis.*

Sr.#	Sequence	NLS signal
1	VangHSP_03	106 - KRR - 108
2	VangHSP_04	239 - SKKRRR - 244
3	VangHSP_07	237 - GRRKR - 241
4	VangHSP_12	221-KEKRKDLEEALTKKRRQI-238

Table 6. Predicted nuclear localization signal (NLS) in *Vigna radiata.*

Sr.#	Sequence	NLS signals
1	VradHSP_03	239 - SKKRRR - 244
2	VradHSP_08	109 - RKKSWVRK - 116

Table 7. *Predicted* nuclear localization signal (NLS) in *Vigna ungucuilata.*

Sr.#	Sequence	NLS signals
1	VunHSP_02	188 - ERERKHLGEKKRR - 200

***Phylogenetic analysis, intron-exon architecture and motif*:** In genome-wide analysis, a total of 15 conserved motifs were identified using MEME tool and conserved domains were identified using ScanProsite and CD search tool in each *Vigna* species. The pattern of NLS and protein cellular localization was predicted using NLStradamus and DeepLoc. 2.5, respectively. A phylogenetic tree was constructed to evaluate the evolutionary relationship among three *Vigna* species from their retrieved sequences of HSP proteins using MEGA7 software. This phylogenetic tree was compared to conserved motif distribution in HSP proteins and the exon-intron architecture of HSP genes. The comparative study on *V. angularis* was carried out using their retrieved HSP sequences. The 17 VangHSP sequences were employed to generate a neighbor joining tree. Their phylogenetic relationship revealed two major clusters (cluster I and II). Among them, cluster I was the largest cluster having 12 HSP proteins. Their exon-intron architecture confirmed all HSP proteins were intronless. Further, a total of 12 conserved protein motifs were found in cluster I (motif 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 14 & 15). VangHSP_01, VangHSP_07, VangHSP_10, and VangHSP_16 (9 motifs) had maximum number of motif among all other member of this cluster analysis. While, the cluster II was comprised of five HSP proteins. Their gene structure display study revealed all HSP proteins were intronless. A total of 12 conserved protein motifs were found in cluster II (motif 1, 2, 3, 4, 5, 6, 8, 10, 11, 12, 14 & 15). VangHSP_05 (10 motifs) and VangHSP_06 (10 motifs) had maximum number of motif among all other member of this cluster analysis (Figure 1).

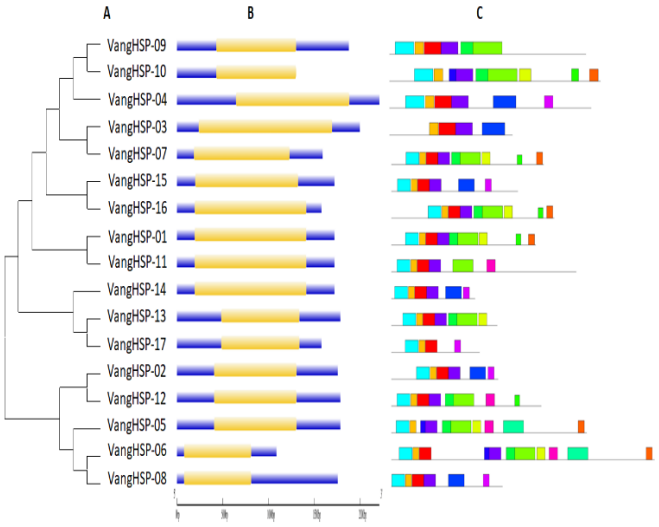

Figure 1. Phylogenetic analysis of *V. angularis*, gene structure display and conserved motif distribution (A) Phylogenetic tree showed two major HSP clusters. (B) Gene structure analysis indicated that most of the HSP genes from *Vigna* species were intronless. (C) Several conserved motifs of HSP were shown. These motifs were coded by 15 different colors.

The comparative study on *V. radiata* was carried out using their retrieved HSP sequences. The 22 VradHSP sequences were employed to generate a neighbor joining tree. Their phylogenetic relationship revealed two major clusters (cluster I and II). Among them, the cluster I was comprised of fourteen HSP proteins. Their gene structure display study revealed all HSP proteins were intronless. A total of 15 conserved protein motifs were found in cluster 1 (motif 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 & 15). VradHSP_07had maximum number of motif (11 motifs) among all other member of this cluster analysis. While, cluster II comprised of 12HSP proteins. Their exon-intron architecture confirmed all HSP proteins were intronless. Further, a total of 15 conserved protein motifs were found in cluster II (motif 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14 & 15). VradHSP_16 had maximum number of motif among all other member of this cluster analysis (10 motifs) (Fig. 2).

The comparative study on *V. ungucuilata* was carried out using their retrieved HSP sequences. The 4 VunHSP sequences were employed to generate a neighbor joining tree. Their phylogenetic relationship revealed one major cluster. Their gene structure display study revealed all HSP proteins were intronless. A total of 14 conserved protein motifs were found (motif 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14 & 15). VunHSP_01had maximum number of motif (13motifs) among all other member of this cluster analysis (Fig. 3).

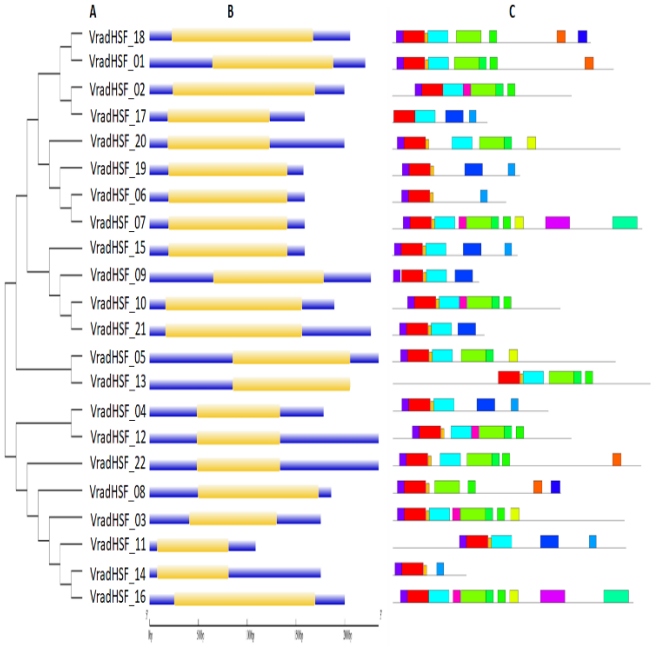

Figure 2. Phylogenetic analysis of *V radiata*, gene structure display and conserved motif distribution (A) Phylogenetic tree showed two major HSP clusters. (B) Gene structure analysis indicated that most of the HSP genes from *Vigna* species were intronless. (C) Several conserved motifs of HSP were shown. These motifs were coded by 15 different colors.

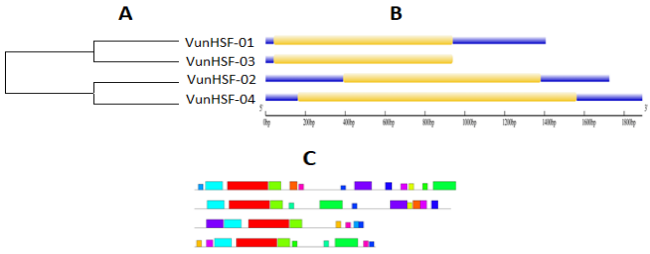

Figure 3. Phylogenetic analysis of *V. ungucuilata*, gene structure display and conserved motif distribution (A) Phylogenetic tree showed two major HSP clusters. (B) Gene structure analysis indicated that most of the HSP genes from *Vigna* species were intronless. (C) Several conserved motifs of HSP were shown. These motifs were coded by 15 different colors.

DISCUSSION

Vigna species are widely recognized legumes that are cultivated all over the world for a variety of uses, including food, fiber, animal feed, and energy generation. It is imperative to understand the factors that control growth and stress resilience for the purpose of developing *Vigna* cultivars that are sustainable. The ultimate goal of this research is to increase agricultural output in the face of challenging environmental circumstances by characterizing HSP protein family in three *Vigna* species (Haider *et al.*, 2021b). A thorough understanding of the molecular processes behind plant stress resistance is essential for ensuring food security in the face of climate change and an expanding global population. In order to overcome the obstacles brought on by changing climatic circumstances, our research aims to disentangle these systems and access genetic resources associated with higher agricultural yields (Haider *et al.*, 2021a; Ijaz *et al.*, 2023a).

Heat Shock Factors (HSFs) have been systematically determined in a range of plant species, including 21 HSFs in *Arabidopsis thaliana* (Nover *et al.*, 2001), 25 each in *Zea mays* (Haider *et al.*, 2021b) and *Oryza sativa* (Shamshad et al., 2023), 20 in soybean (Lopes-Caitar *et al.*, 2013), and 61 HSFs in *Triticum aestivum* (Nagaraju et al., 2015). In order to conduct focused research and develop crop improvement techniques, it is important to have a detailed understanding of the complex regulatory network that governs plants' heat shock response. This is shown by the variation in HSF presence across various plant genomes. However, in the current study, a total of 43 non-redundant HSP proteins were identified in three Vigna species: 17 in *V. angularis*, 22 in *V. radiata*, and 4 in *V. ungucuilata*.

Moreover, 15 conserved motifs with distinct architecture were identified in the *Vigna* species using the MEME search tool (Bailey *et al.*, 2009). This number of motifs in the HSP family has been validated by previous research studies that also found fifteen conserved motifs in various crops (Habib., 2022; Diogo-Jr *et al.*, 2023; Shamshad *et al.*, 2023).The physicochemical properties of HSP proteins of three *Vigna* revealed that Leucine (Leu) was found as a highly abundant amino acid followed by glutamic acid (Glu), valine (Val), aspartic acid (Asp), and arginine (Arg) and threonine (Thr) in *V. angularis*, arginine (Arg) was found as highly abundant amino acid followed by aspartic acid (Asp), Glutamine (Glu), serine (Ser), and valine (Val) in *V. radiata*, and serine (Ser) was found as highly abundant amino acid followed by leucine (Leu) and thrionine (Thr), asparginine (Asn), and glutamine (Glu) in *V. ungucuilata*. These amino acids play an important role in various biological and physiological processes under stress conditions in plants (Trovato *et al.*, 2021).

The theoretical isoelectric point ranged from 4.93 to 9.50 in *V. angularis*, 4.69 to 9.30 in *V.radiata*, and 5.96 to 8.89 in *V. ungucuilata*. Most of the HSP proteins were in the acidic range of pI value and previous studies also validated that most of the HSP proteins were found to be acidic (Haider *et al.*, 2021b). The instability index is a stability measure of proteins and if the instability index is less than 40 then protein is considered to be stable (Ijaz *et al.*, 2023a). The instability index (II) was 39.27 and 32.93 for VangHSP_9 and VradHSP_9, considered as stable while the rest of the HSP proteins were unstable. In a previous study, most of the HSPs were stable with instability index < 40 (Baloji *et al.*, 2019; Jakhu *et al.*, 2021). However, in *V. ungucuilata*, all HSP proteins were unstable with more than 40 instability index. The aliphatic index ranged from 61.98 to 85.36 in *V. angularis*, 61.41 to 79.69 in *V. radiata* and from 65.43 to 73.62 in *V. ungucuilata*. The aliphatic index of the HSP proteins demonstrated that all proteins are highly thermostable and they can endure adversative ecological circumstances (Panja *et al.*, 2020). Proteins are categorized as stable if they're in vitro stability is greater than a cut-off value of >40 and as unstable otherwise by the instability index (Ijaz *et al.*, 2022) and this was validated by previously reported research that also predicted most of the HSP proteins as thermally stable (Baloji *et al.*, 2019).

Subcellular localization is the location of proteins that are present in plant cells. In general, all HSP proteins are present in the nucleus (Tu *et al.*, 2023). As a result, three *Vigna* species HSP subcellular locations have also been predicted that all HSPs were found in the nucleus. Moreover, 4 NLSs were predicted in *V. angularis* HSPs, two NSLs were predicted in *V. radiate*, in contrast, only one NLS was predicted in *V. ungucuilata*. Nuclear localization signals are imperative for predicting nuclear import and this is directed by the strength of nuclear localization signals that identifies that intracellular dispersal in plant cells (Lin *et al.*, 2011). Moreover, a phylogenetic tree was constructed to evaluate the evolutionary relationship among three *Vigna* species from their retrieved sequences of HSP proteins using MEGA7 software. This phylogenetic tree was compared to conserved motif distribution in HSP proteins and the exon-intron architecture of HSP genes. The phylogenetic relationship of all three Vigna species revealed two major clusters (cluster I and II). Further, their exon-intron architecture confirmed all HSP proteins were intronless and these findings were in consistent with previous studies that also found that all tendem duplicated HSP genes were intronless in potato (Zhao *et al.*, 2018).

***Conclusion*:** In conclusion, the Heat Shock Protein (HSP) family in Vigna species has been the subject of a genome-wide in silico research that has sheded important insights into the molecular processes driving stress response in these plants. It is possible to gain a better knowledge of the distribution, organization, and potential regulatory functions of HSP genes in dealing with different environmental stressors by methodically analyzing the genomic architecture of these genes. These results establish the foundation for additional experimental studies aimed at verifying the anticipated roles of certain HSPs and clarifying their roles in stress resilience in *Vigna* species. In the end, our work advances the more general objective of creating robust crop types that can flourish under harsh environmental circumstances.

***Funding*:** N/A

***Conflict of Interest*:** We clarify that the submitted manuscript is our original research work and has not been published previously. Additionally, there is no competing interest.

Ethical statement: This article does not contain any research with human participants or animal performed by any of the authors.

***Availability of data and material*:** We declare that the submitted research article is our own work which has not been published before and is not currently considered for elsewhere publication.

***Acknowledgement*:** This research work was conducted and supported by the Molecular Biology of Plant Disease Resistance Laboratory, Centre of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture, Faisalabad.

***Code availability*:** N/A

***Consent for publication*:** All authors have given the consent to publish this manuscript in Phytopathogenomics and Disease Control (PDC).

REFERENCES

Ahanger, M.A., N.A. Akram, M. Ashraf, M.N. Alyemeni, L. Wijaya and P. Ahmad. 2017. Plant responses to environmental stresses—from gene to biotechnology.AoB Plants 9:plx025.

Atteh, A. and A. Adeyeye. 2022. Effect of low gamma irradiation on the germination and morphological characteristics of broad beans (Vicia faba L.), mung beans (Vigna radiata L.), and peas (Pisum sativum L.) seedlings. Natural resources 13:105-125.

Bailey, T.L., M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi and J.L. Ren. 2009. WW:, Noble WS MEME Suite: Tools for motif discovery and searching. Nucleic Acids Research 37:202-208.

Baloji, G., S. Pasham, V. Mahankali, M. Garladinne and S. Ankanagari. 2019. Insights from the molecular docking analysis of phytohormone reveal brassinolide interaction

with HSC70 from Pennisetum glaucum. Bioinformation 15:131.

Bhardwaj, R., J.K. Lone, R. Pandey, N. Mondal, R. Dhandapani, S.K. Meena and S. Khan. 2023. Insights into morphological and physio-biochemical adaptive responses in mungbean (Vigna radiata L.) under heat stress. Front. Genet 14:1206451.

Chen, X., Z. Wang, R. Tang, L. Wang, C. Chen and Z. Ren. 2021. Genome-wide identification and expression analysis of Hsf and Hsp gene families in cucumber (Cucumis sativus L.). Plant Growth Regul 95:223-239.

Diogo-Jr, R., E.V. de Resende Von Pinho, R.T. Pinto, L. Zhang, J.A. Condori-Apfata, P.A. Pereira and D.R. Vilela. 2023. Maize heat shock proteins-prospection, validation, categorization and in silico analysis of the different ZmHSP families. Stress Biologly 3:37.

Habib, Z. n.d. DNA-binding One Zinc Finger (DOF) transcription factors in Vigna species: A review.

Haider, S., J. Iqbal, S. Naseer, T. Yaseen, M. Shaukat, H. Bibi, Y. Ahmad, H. Daud, N.L. Abbasi and T. Mahmood. 2021a. Molecular mechanisms of plant tolerance to heat stress: current landscape and future perspectives. Plant Cell Rep.Pp.1-25.

Haider, S., S. Rehman, Y. Ahmad, A. Raza, J. Tabassum, T. Javed, H.S. Osman and T. Mahmood. 2021b. In silico characterization and expression profiles of heat shock transcription factors (HSFs) in maize (Zea mays L.). Agronomy 11:2335.

Ijaz, S., I.U. Haq, I.A. Khan, H.M. Ali, S. Kaur and H.A. Razzaq. 2022. Identification of resistance gene analogs of the NBS-LRR family through transcriptome probing and in silico prediction of the expressome of Dalbergia sissoo under dieback disease stress. Front. Genet 2864.

Ijaz, S., I.U. Haq, R. Malik, G. Nadeem, H.M. Ali and S. Kaur. 2023a. In silico characterization of differentially expressed short-read nucleotide sequences identified in dieback stress-induced transcriptomic analysis reveals their role as antimicrobial peptides. *Front.* Plant Sciences 14:1168221.

Ijaz, S., I.U. Haq, H.A. Razzaq, B. Nasir, H.M. Ali and S. Kaur. 2023b. In silico structural-functional characterization of three differentially expressed resistance gene analogs identified in Dalbergia sissoo against dieback disease reveals their role in immune response regulation. Front. Plant Sciences 14.

Jakhu, P., P. Sharma, I.S. Yadav, P. Kaur, S. Kaur, P. Chhuneja and K. Singh. 2021. Cloning, expression analysis and In silico characterization of HSP101: a potential player conferring heat stress in Aegilops speltoides (Tausch) Gren. Physiology and Molecular Biology of Plants 27:1205-1218.

Juneja, S., R. Saini, A. Adhikary, R. Yadav, S.A. Khan, H. Nayyar and S. Kumar. 2023. Drought priming evokes essential regulation of Hsp gene families, Hsfs and their related miRNAs and induces heat stress tolerance in chickpea. Plant Stress 10:100189.

Karami, Z., C. Butkinaree, Y. Yingchutrakul, N. Simanon and K. Duangmal. 2022. Comparative study on structural, biological and functional activities of hydrolysates from Adzuki bean (Vigna angularis) and mung bean (Vigna radiata) protein concentrates using Alcalase and Flavourzyme. Food Research International 161:111797.

Lin, Y.-X., H.-Y. Jiang, Z.-X. Chu, X.-L. Tang, S.-W. Zhu and B.-J. Cheng. 2011. Genome-wide identification, classification and analysis of heat shock transcription factor family in maize. BMC Genomics 12:1-14.

Lopes-Caitar, V.S., M.C.C.G. de Carvalho, L.M. Darben, M.K. Kuwahara, A.L. Nepomuceno, W.P. Dias, R. V Abdelnoor and F.C. Marcelino-Guimarães. 2013. Genome-wide analysis of the Hsp 20 gene family in soybean: comprehensive sequence, genomic organization and expression profile analysis under abiotic and biotic stresses. BMC Genomics 14:1-17.

Mondal, S., S. Karmakar, D. Panda, K. Pramanik, B. Bose and R.K. Singhal. 2023. Crucial plant processes under heat stress and tolerance through heat shock proteins. Plant Stress 100227.

Nagaraju, M., P. Sudhakar Reddy, S. Anil Kumar, R. K Srivastava, P.B. Kavi Kishor and D.M. Rao. 2015. Genome-wide scanning and characterization of Sorghum bicolor L. heat shock transcription factors. *Curr.* Genomics 16:279-291.

Nover, L., K. Bharti, P. Döring, S.K. Mishra, A. Ganguli and K.-D. Scharf. 2001. Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? Cell Stress Chaperones 6:177.

Panja, A.S., S. Maiti and B. Bandyopadhyay. 2020. Protein stability governed by its structural plasticity is inferred by physicochemical factors and salt bridges. Scientific Reports 10:1822.

Ren, Y., R. Ma, M. Xie, Y. Fan, L. Feng, L. Chen, H. Yang, X. Wei, X. Wang and K. Liu. 2023. Genome-wide identification, phylogenetic and expression pattern analysis of HSF family genes in the Rye (Secale cereale L.). BMC Plant Biology 23:441.

Shamshad, A., M. Rashid and Q. uz Zaman. 2023. In-silico analysis of heat shock transcription factor (OsHSF) gene family in rice (Oryza sativa L.). BMC Plant Biology .pp. 23:395.

Singh, P., B. Pandey, A. Pratap, U. Gyaneshwari, R.M. Nair, A.K. Mishra and C.M. Singh. 2022. Genetic and genomics resources of cross-species vigna gene pools for improving biotic stress resistance in mungbean (vigna radiata l. Wilczek). Agronomy 12:3000.

Trovato, M., D. Funck, G. Forlani, S. Okumoto and R. Amir. 2021. Amino acids in plants: Regulation and functions in development and stress defense.

Tu, J., M. Abid, J. Luo, Y. Zhang, E. Yang, X. Cai, P. Gao, H. Huang and Z. Wang. 2023. Genome-wide identification of the heat shock transcription factor gene family in two kiwifruit species. Frontiers in Plant Science 14:1-14

Tutlani, A., P.J. Banshidhar and H.S. Janeja. 2022. Abiotic and biotic stresses and their effect on Vigna radiata L. Journal of Pharmaceutical Innovation 11:230-237.

Wang, Y., X. Yao, H. Shen, R. Zhao, Z. Li, X. Shen, F. Wang, K. Chen, Y. Zhou and B. Li. 2022. Nutritional composition, efficacy, and processing of Vigna angularis (Adzuki Bean) for the human diet: An overview. Molecules 27:6079.

Wu, J., T. Gao, J. Hu, L. Zhao, C. Yu and F. Ma. 2022. Research advances in function and regulation mechanisms of plant small heat shock proteins (sHSPs) under environmental stresses. Science of the Total Environment 825:154054.

Zhang, H., Z. Yu, X. Yao, J. Chen, X. Chen, H. Zhou, Y. Lou, F. Ming and Y. Jin. 2021. Genome-wide identification and characterization of small auxin-up RNA (SAUR) gene family in plants: evolution and expression profiles during normal growth and stress response. BMC Plant Biology 21:1-14.

Zhao, P., D. Wang, R. Wang, N. Kong, C. Zhang, C. Yang, W. Wu, H. Ma and Q. Chen. 2018. Genome-wide analysis of the potato Hsp20 gene family: identification, genomic organization and expression profiles in response to heat stress. BMC Genomics 19:1-13.

Zhou, Y., Y. Wang, F. Xu, C. Song, X. Yang, Z. Zhang, M. Yi, N. Ma, X. Zhou and J. He. 2022. Small HSPs play an important role in crosstalk between HSF-HSP and ROS pathways in heat stress response through transcriptomic analysis in lilies (Lilium longiflorum). BMC Plant Biolology 22:202.